

Easy and efficient data management in Stata

Bill Rising

StataCorp LP

Stata Conference DC 09

Outline I

- 1 Introduction
 - Goals
- 2 GUI to Do
 - Working Reproducibly
 - Data Management Notes
- 3 Conclusion
 - Conclusion

Goals

- Learn to use the new GUI tools in Stata
 - Data Editor
 - Variables Manager
 - Do-file Editor
- Learn to work in a reproducible fashion with the GUI

Preparing

- To make commands reproducible, we need to
 - Start a command log
 - Here: `testdrive.do`
 - Specify the `.do` extension explicitly
 - Open the dataset
 - Open the Data Editor
 - Set the sort order to something reproducible
 - Best to sort on a primary key
- Only then can we start editing

From here on

- To quote Enders (2007)
“point click point point click click”
 - Extended by Wiggins (2007)
“point click point point click click **click**”

Data Editor notes

- The Data Editor is no longer modal
 - May be open while working
- Can edit or enter dates as dates
- Can edit date formats by example
- Can filter observations or variables without jumping in and out of the editor
- Can save and manage snapshots
- Can work with variable properties

Data Editor extras

- Can jump to specific cell with *varname obs #*
 - No need for brackets
 - Can use variable abbreviations

Variables Manager notes

- Can manage all that has to do with structure
- Can work with variable properties
- Can manage notes

Variables Manager extras

- Can change order of columns
- Can group variables by column
- Can sort variable list by column
 - Useful for finding blank variable labels
 - Useful for

Data Editor notes

- Can fold code between braces
 - Neat trick 1: commented blocks
- Neat trick 2: Put code in blocks, ctrl-b ctrl-d will run block

merge changes

- Syntax is greatly simplified
- Produces table for checking automatically
- Has some convenience options

append changes

- Can append many files in one command
- Can track which observations came from which file
- Can append to an empty dataset
 - Great trick: `append using `: dir . files "*.dta"'`

Conclusion

- Data Editor now useful
- Variables Manager nice for big datasets
- Do-file Editor has some nice tricks